

## One more time: bibliometric analysis of scientific output remains complicated

T. Opthof · A. A. M. Wilde

Published online: 20 July 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Dear Sir,

In the May issue of the Netherlands Heart Journal, we published an update [1] of a previous bibliometric analysis of the scientific work of Dutch professors in clinical cardiology [2]. The paper was accompanied by an editorial [3] in the same issue and another one in the subsequent June issue [4].

Van der Wall [3] summarises our paper and emphasises that we have stated ‘...that citation analysis should always be applied with great care in science policy.’ We indeed concluded that given the inhomogeneity in citation characteristics of the scientific output of an -at first sight-homogeneous group of clinical cardiologists, such analyses lack sufficient validation for application in a more complex network as a university medical centre. In his 2005 paper [5], Hirsch indeed proposed his well-known *h*-index for quality assessment of scientific output. At that time he emphasised that Nobel prize winners had *h*-indices of 30 and higher. In addition he mentioned that a *h*-index of 18 might be reasonable as an equivalent for a full-professorship, whereas a *h*-index of 10–12 might be reasonable to obtain ‘tenure’. It is important to underscore that Hirsch’s statements were restricted to the arena of physics, where *h*-indices are much lower than in the life sciences, as pointed out previously by us [2] and others [6].

Professor Doevedans’ contribution [4] focuses on the competitive aspects of science in relation with its financial system rather than on our paper [1]. We fully agree with Doevedans that the credibility of a funding system is very important. A system primarily based on grants heavily leans on the integrity of committees and individuals deciding on the fate of research proposals. With respect to this the literature is not very reassuring to put it mildly [7–9]. There is a trend towards decision-making at an early stage (within academic medical centres; i.e. *before* actual submission) and to decision-making *without* peer review after submission. Despite the understandable efforts to save time and energy both on the part of applicants and those in control of the system, we share the concerns of Professor Doevedans [4]. Therefore we can imagine that bibliometric analysis can be perceived as part of an armamentarium in a tombola with unequal chances and is therefore received with distrust.

Having said that, we must take away from three of Doevedans’ suggestions. 1) Authors with an uncommon name had a disadvantage of up to 20%, whereas data of authors with ‘a more common name’ were polluted. It is suggested that we allocated citations or papers to professors that they did not receive, respectively write. In case of doubt, our data were checked on a *per author/per paper* base. Therefore our data stand! However, in general, Professor Doevedans is correct in emphasising that when an assessment is made by organisations/institutions that do not know the individuals/work under assessment, there is an increased risk for errors. This implies that an adversary approach is always wise. But again, this is not the case in our study. 2) Doevedans’ claim that the use of alternative databases as *Scopus* (or e.g. *Google Scholar*) instead of Thomson Reuters’ *Web of Science (WoS)* would lead to different rankings is not substantiated with data. However,

---

T. Opthof (✉) · A. A. M. Wilde  
Department of Clinical and Experimental Cardiology,  
Heart Failure Research Center, Academic Medical Center,  
Meibergdreef 9, Room K2-105,  
1105 AZ Amsterdam, the Netherlands  
e-mail: t.opthof@inter.nl.net

T. Opthof  
Department of Medical Physiology,  
University Medical Center Utrecht,  
Utrecht, the Netherlands

it would certainly be interesting to see the effect on rankings. Some databases index more journals than others and with different time lags as well. To us it seems improbable that individuals in any top-10 based on the Thomson Reuters' database would not score high in the other databases as well. But the onus to show this is not on us, but on those who criticise our study. 3) Doevendans states that 'although the paper is accepted by a scientific journal, the reproducibility of the data has not been established.' This is simply not true because a second count of several authors, using the same database (i.e. *WoS*), revealed identical results. Within one and the same database, the data are reproducible.

It is a pity that Professor Doevendans focuses on differences between the databases without an effort to either measure or explain these. In addition, there is some juggling with the concept of 'reproducibility'. This tends to obscure the quintessence of our paper: large differences in citations and also in *h*-indices occur within a relatively homogeneous group of clinical cardiologists. These differences cannot be interpreted as differences in scientific quality, because we have shown that there are large differences in citation frequency, e.g. between 'sub-subfields' as 'Marfan syndrome' and 'Brugada syndrome', but also between 'subfields' as 'congenital heart disease' and 'arrhythmias', just to mention a few. Thus, this type of citation analysis at the personal level should be discouraged, because it can damage the scientific status of individuals. It goes without saying that this advice is even more urgent when it comes to comparison of different specialisations within university medical centres, where these are meant to corroborate each other.

We finish by summarising three important parameters that determine the total of citations of a scientist. First is the network. How many co-authors (and grants...) were involved? Second is the 'citation culture'. In medicine, papers have many more references than in mathematics or computer science. We would not want to see a mathematics faculty closed down because the *h*-indices of its scientists are considered too low (compared with the medicine faculty). Third is the number of scientists publishing in a field. If an author publishes a paper and there are 10 other scientists active and publishing in the same field, whose papers all cite the work, the author under assessment obtains 'only' 10 citations, despite the fact that 100% of the

citing authors have cited the work. If there are 100 other scientists of which 50 cite the work, the author under assessment obtains 50 citations (five times as many) although 'only' 50% of the citing authors have cited the work. These three issues, in particular the last one, have thus far not been fully addressed in the specialised literature in such a way that consensus resulted.

There is an additional concern. When an identical paper is simultaneously published in two or more journals, the citations obtained are not identical, but strongly correlated with the impact factor of the publishing journal [10]. The difference can be substantial. How then can citations be taken as a parameter of scientific quality?

For all these reasons we concluded that 'citation analysis should be applied with great care in science policy'. And not because our data are incorrect or not reproducible.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Opthof T, Wilde AAM. Bibliometric data in clinical cardiology revisited. The case of 37 Dutch professors. *Neth Heart J*. 2011;19:246–55.
2. Opthof T, Wilde AAM. The Hirsch-index: a simple, new tool for the assessment of scientific output of individual scientists. The case of Dutch professors in clinical cardiology. *Neth Heart J*. 2009;17:145–54.
3. Van der Wall EE. The Hirsch index: more than a number? *Neth Heart J*. 2011;19:209–10.
4. Doevendans PA. To publish or not to publish. *Neth Heart J*. 2011;19:271–2.
5. Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA*. 2005;102:16569–72.
6. Ball P. Index aims for fair ranking of scientists. *Nature*. 2005;436:900.
7. Langfeldt L. Judging quality. Expert panels evaluating research: decision-making and sources of bias. *Research Eval*. 2004;13:51–62.
8. Moed HF. Citation analysis in research evaluation. Dordrecht: Springer; 2005. p. 229–57.
9. Bommann L, Leydesdorff L, Van den Besselaar P. A meta-evaluation of scientific research proposals: different ways of comparing rejected to awarded applications. *J Informetr*. 2010;4:211–20.
10. Opthof T. Sense and nonsense about the impact factor. *Cardiovasc Res*. 1997;33:1–7.